In [ ]:
```
import pandas as pd
import matplotlib.pyplot as plt
```

In [ ]:
```
df=pd.read_csv(r"C:\Users\advik\OneDrive\Desktop\experiment4.csv")
```

In [ ]:
```
print(df)
```

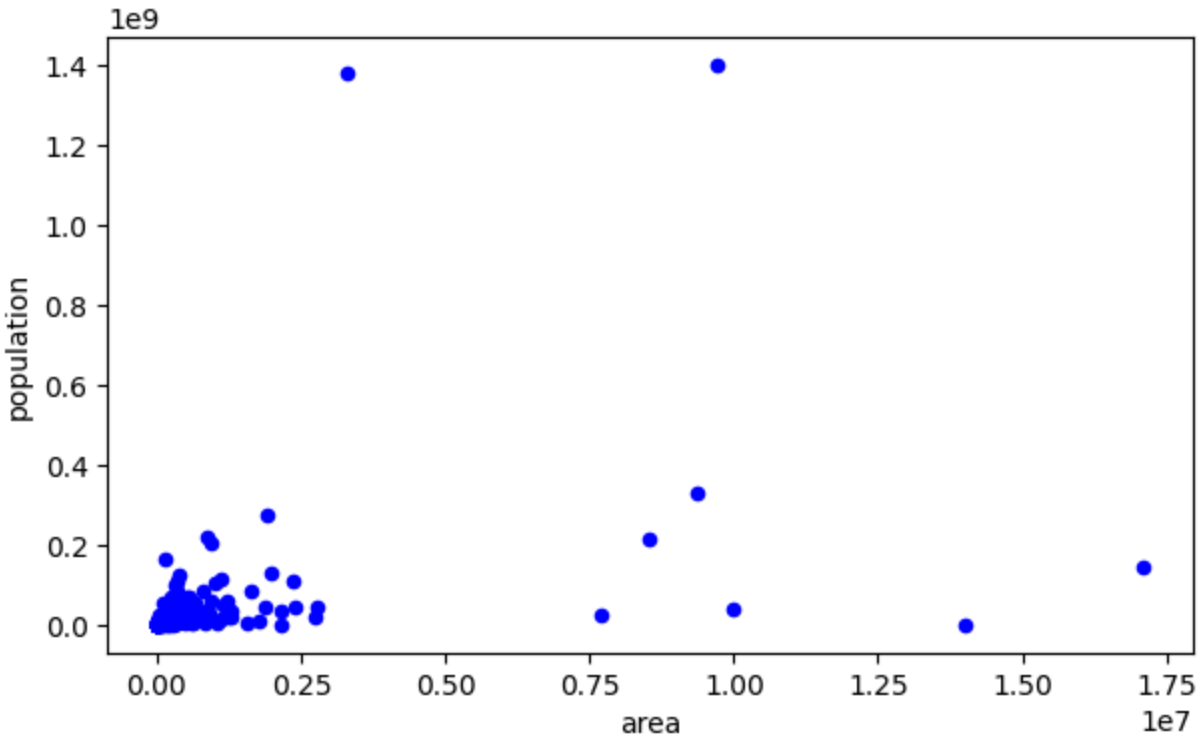
```
              name            capital  population      area     region
0    South Georgia  King Edward Point          30    3903.0  Antarctic
1          Grenada       St. George's      112519     344.0   Americas
2      Switzerland               Bern     8654622   41284.0     Europe
3     Sierra Leone           Freetown     7976985   71740.0     Africa
4          Hungary           Budapest     9749763   93028.0     Europe
..             ...                ...         ...       ...        ...
245        Belgium           Brussels    11555997   30528.0     Europe
246         Israel          Jerusalem     9216900   20770.0       Asia
247    New Zealand         Wellington     5084300  270467.0    Oceania
248      Nicaragua            Managua     6624554  130373.0   Americas
249       Anguilla         The Valley       13452      91.0   Americas

[250 rows x 5 columns]
```

Scatter plot

In [ ]:
```
df.plot(x='area', y='population', kind='scatter', color='blue', figsize=(7,4))
#A scatter plot is a type of graph that displays individual data points as dot
#  The x-axis typically represents one variable, and the y-axis represents the
```

Out[ ]:
```
<Axes: xlabel='area', ylabel='population'>
```


1e9
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
population
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
1e7
area

Here is a wide range of both area and popultaion sizes .

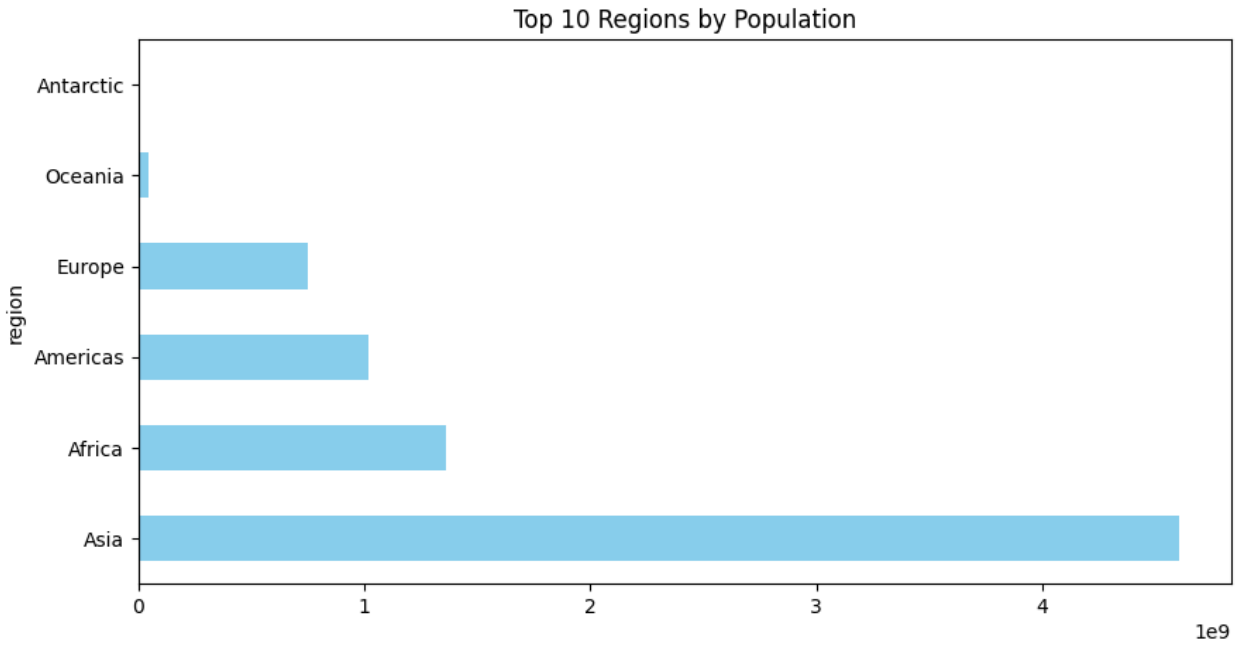
The majority of countries have smaller areas and populations, with a few outliers having very large areas and / or populations.

There does appear to be a positive correlation between areas and population , which makes sense intuitively - larger countries tend to have larger populations. However , this relationship is not linear , and theres a lot of variability.

Horizontal Bar plot

In [ ]:
```
df.groupby('region')['population'].sum().sort_values(ascending=False).head(10)
    kind="barh", figsize=(10, 5), color="skyblue", title="Top 10 Regions by Po
# bar plot (or bar chart) is a type of graph used to display and compare the f
# It consists of rectangular bars where the length of each bar is proportional
```

Out[ ]:
```
<Axes: title={'center': 'Top 10 Regions by Population'}, ylabel='region'>
```



The significantly larger population of Asia compared to other regions.

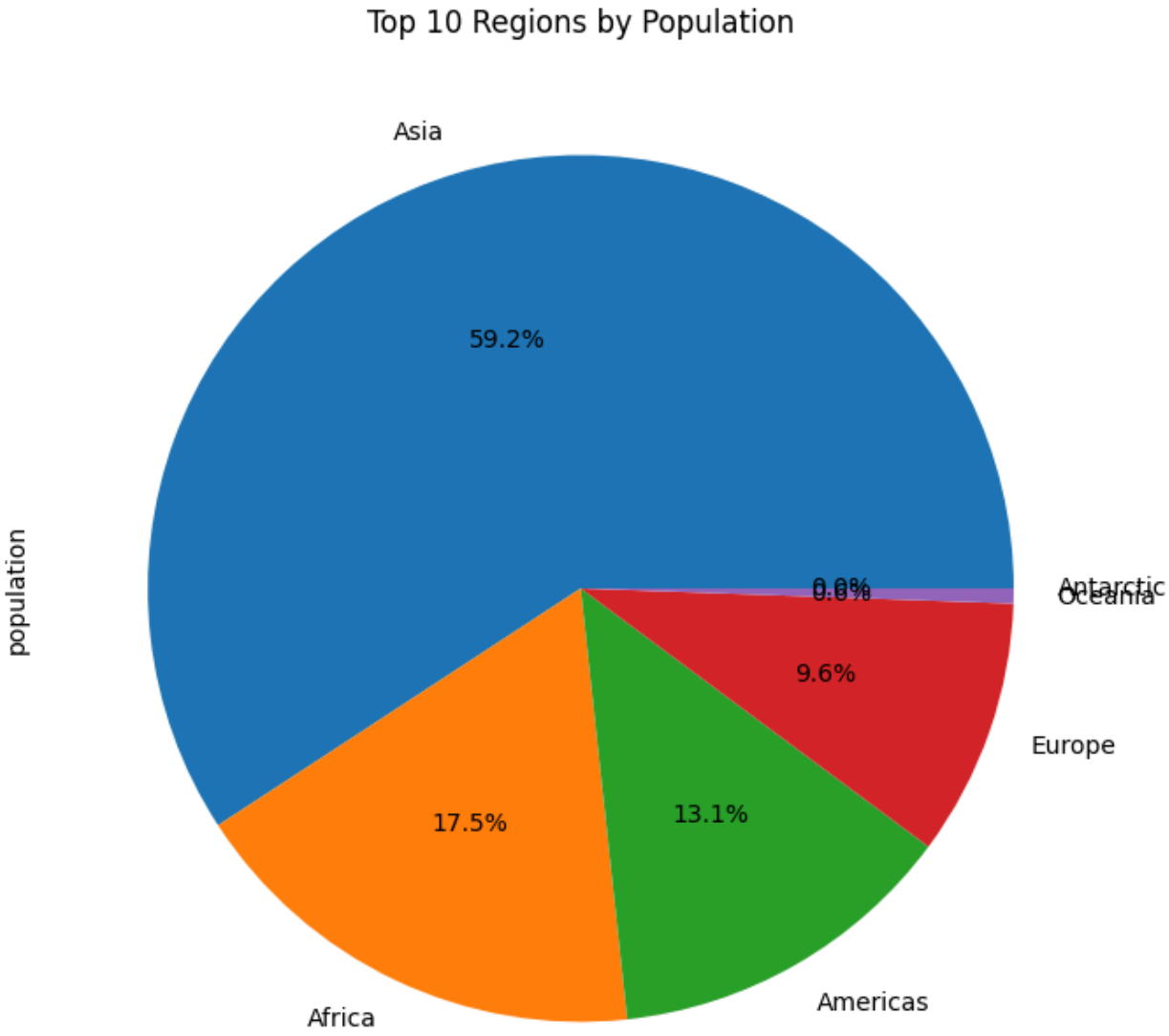
The relative distribution of populations across the Americas, Europe, and Africa.

The comparatively small populations in Oceania.

Pie plot

In [ ]:
```
df.groupby('region')['population'].sum().sort_values(ascending=False).head(10)
    kind="pie", autopct='%1.1f%%', figsize=(8, 8), title="Top 10 Regions by Po
#A pie chart (or pie plot) is a circular chart divided into sectors, each repr
#  It's commonly used to show how different parts make up a whole, especially
```

Out[ ]:
```
<Axes: title={'center': 'Top 10 Regions by Population'}, ylabel='population'>
```


Top 10 Regions by Population
Asia
59.2%
population
0.0%
Antarctic
Oceania
9.6%
Europe
17.5%
13.1%
Africa
Americas

Asia acconts for the highest total population , followed by the Americas , and then Africa . Europe and Oceania have significantly similar populations compared to the other regions.

pie chart effectively visualize the proportion of the population that each region represents making it easy to compare the relative sizes of each regions population.